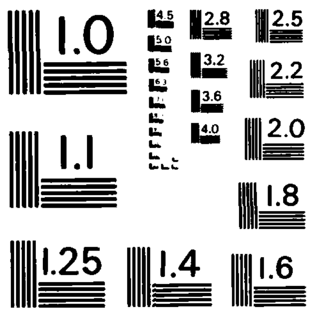
AD-A132 863 MORE RESULTS ON THE CONVERGENCE OF ITERATIVE METHODS FOR THE SYMMETRIC LI..(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER J PANG AUG 83 MRC-TSR-2552 1/1

UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL

END
DATE
FILMED
10 83
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A133 963

MRC Technical Summary Report #2552

MORE RESULTS ON THE CONVERGENCE OF ITERATIVE METHODS FOR THE SYMMETRIC LINEAR COMPLEMENTARITY PROBLEM

Jong-Shi Pang

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

August 1983

(Received June 14, 1983)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

MORE RESULTS ON THE CONVERGENCE OF ITERATIVE METHODS FOR THE SYMMETRIC LINEAR COMPLEMENTARITY PROBLEM[1,2]

Jong-Shi Pang

Technical Summary Report #2552
August 1983

ABSTRACT

In an earlier paper, the author has given some necessary and sufficient conditions for the convergence of iterative methods for solving the linear complementarity problem. These conditions may be viewed as global in the sense that they apply to the methods regardless of the constant vector in the linear complementarity problem. More precisely, the conditions characterize a certain class of matrices for which the iterative methods will converge, in a certain sense, to a solution of the linear complementarity problem for all constant vectors. In this paper, we improve on our previous results and establish necessary and sufficient conditions for the convergence of iterative methods for solving each individual linear complementarity problem with a fixed constant vector. Unlike t.e earlier paper, our present analysis applies only to the symmetric linear complementarity problem. Various applications to a strictly convex quadratic program are also given.

AMS (MOS) Subject Classifications: 90C20, 90C33

Key Words: Convergence, iterative methods, linear complementarity problem, quadratic programming, matrix splitting.

Work Unit Number 5 - Optimization and Large Scale Systems

---

[1]This research was based on work supported by the National Science Foundation under Grant No. ECS-8114571 and was completed while the author was visiting the Mathematics Research Center at the University of Wisconsin-Madison.

[2]The author gratefully acknowledges several stimulating conversations with Professor Olvi Mangasarian.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

-A-

SIGNIFICANCE AND EXPLANATION

Iterative methods have been found very useful for solving large-scale linear complementarity and quadratic programming problems. There are many conditions proven to be sufficient for such methods to converge. However, there are very few conditions that are known to be necessary for convergence. Necessary conditions are useful because they identify the underlying limitation of the methods. In an earlier paper, we have been able to derive some necessary and sufficient conditions for the convergence of a large class of iterative methods. In the present paper, we improve on our earlier results and establish among other things, the convergence of many iterative methods under the weakest possible conditions.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

MORE RESULTS ON THE CONVERGENCE OF ITERATIVE METHODS FOR THE SYMMETRIC LINEAR COMPLEMENTARITY PROBLEM[1,2]

Jong-Shi Pang

1. INTRODUCTION

In an earlier paper (Ref. 1), we have established some necessary and sufficient conditions for the convergence of iterative methods for solving the linear complementarity problem. These conditions may be viewed as global in the sense that they apply to the methods regardless of the constant vector in the linear complementarity problem (LCP). Specifically, consider the LCP (q,M):

$$q + Mx \geqq 0, \ x \geqq 0 \quad \text{and} \quad x^T(q+Mx) = 0 \ ,$$

where $q \in R^n$ and $M \in R^{n\times n}$ are given and $x \in R^n$. Let (B,C) be a Q-splitting of the matrix M, i.e. $M = B+C$ and B is a Q-matrix (the LCP(q,B) has a solution of all vectors q). Let E be a nonnegative diagonal matrix with $E_{ii} < 1$. Define the point-to-set algorithmic map A as follows: for all vectors x,

$$A(x) = \text{solution set of the LCP}(q+Cx,\ B,\ Ex) \ .$$

The latter LCP(r,B,s) is to find y so that

$$r+By \geqq 0, \ y \geqq s \quad \text{and} \quad (y-s)^T(r+By) = 0 \ .$$

Obviously, under the translation of variables $x = y-s$, the LCP(r,B,s) can be converted into the LCP(r+Bs,B). Since B is a Q-matrix, the set A(x) is

---

[1]This research was based on work supported by the National Science Foundation under Grant No. ECS-8114571 and was completed while the author was visiting the Mathematics Research Center at the University of Wisconsin-Madison.

[2]The author gratefully acknowledges several stimulating conversations with Professor Olvi Mangasarian.

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

nonempty for all vectors x. Moreover, a vector x^* solves the LCP(q,M) if and only if it is a fixed point of the map A, i.e. $x^* \in A(x^*)$.

Given the diagonal matrix E and the Q-splitting (B,C) of the matrix M, we define an iterative scheme for solving the LCP(q,M): Let $x^0 \geqq 0$ be an arbitrary nonnegative vector. In general, given $x^k \geqq 0$ $(k \geqq 0)$, let x^{k+1} be any vector in the set $A(x^k)$.

If B is a P-matrix, then the set $A(x)$ is a singleton for all x. In this case, each x^{k+1} will be uniquely defined.

As noted in the earlier papers (Refs. 1 and 2), the above fundamental scheme provides a unifying framework for the study of many well-known iterative methods for solving the LCP(q,M). In Ref. 1, necessary and sufficient conditions on the matrix M have been given so that for all vectors q and all starting vectors $x^0 \geqq 0$, each sequence $\{x^k\}$ generated by the basic iterative scheme will "converge" to some solution of the LCP(q,M). The precise characterization of convergence is dependent on the notion of convergence involved, on whether M is symmetric and on the additional property imposed on the splitting (B,C).

A key feature of the analysis in Ref. 1 is that the characterization applies to convergence for all constant vectors q. For all practical purposes, it would be of interest to obtain some characterization of convergence for each individual q. A major objective of this paper is to derive some necessary and sufficient conditions on both M and q so that the basic iterative scheme defined above will "converge" (to some solution of the LCP(q,M)) for all initial vectors $x^0 \geqq 0$. The notion of convergence is a rather delicate one. As it is already evident from the results in Ref. 1, the kind of convergence that one can characterize depends rather crucially on the assumed properties of the matrix M.

There are two main characterizations obtained (Theorems 2.1 and 3.1). Both require that the matrix M be symmetric. The symmetry of M seems rather indispensable for the kind of characterization sought after in this paper. This is because in the asymmetric case, the analysis employed here breaks down. There, a typical argument for convergence is contraction of the iterates in which the role played by q seems quite minimal. As a result, it might be necessary to search for a different kind of characterization in the asymmetric case.

In addition to the two Theorems 2.1 and 3.1, several characterizations for the convergence of various dual iterative methods for solving a strictly convex quadratic program are derived. All the characterizations are obtained in terms of some very minimal requirements on the original problem being solved. For example, in the case of a strictly convex quadratic program, the convergence of the sequence of primal vectors induced by the dual iterative methods is characterized by the feasibility (or equivalently, solvability) of the primal program under absolutely no other conditions.

Closely related to the LCP(q,M) is the quadratic program

$$\min_{x \geqq 0} f(x) := q^T x + \frac{1}{2} x^T M x \quad . \tag{1}$$

If M is symmetric, then a vector x^* solves the LCP(q,M) if and only if it is a stationary point of the above program. The objective function $f(x)$ plays a crucial role in the analysis that follows.

We explain some matrix notations used in the paper. If A is an $n \times m$ matrix, α and β are subsets of $\{1,\ldots,n\}$ and $\{1,\ldots,m\}$ respectively, by $A_{\alpha\beta}$ we denote the submatrix of A whose rows and columns are indexed by α and β respectively. If $\alpha = \{1,\ldots,n\}$, we denote by $A_{\cdot\beta}$ those columns of A indexed by β. Similar definition applies to $A_{\alpha\cdot}$.

2. THE NONDEGENERATE CASE

We divide our analysis into two cases, depending on whether the matrix M is nondegenerate or positive semi-definite. Recall that a matrix M is nondegenerate if all its principal minors are nonzero. A well-known characterization of nondegeneracy in linear complementarity theory is the following (Ref. 3): M is nondegenerate if and only if the LCP(q,M) has a finite number of solutions for all vectors q.

Before stating our main result for the nondegenerate case, we quote the following lemma whose proof can be found in Ref. 1. The splitting (B,C) of the matrix M is regular if $B-C$ is positive definite.

Lemma 2.1. Let (B,C) be a regular splitting of the symmetric matrix M. Then, for any nonnegative diagonal matrix E with $E_{ii} < 1$ for all i and for any nonnegative vector x,

$$f(x) - f(y) \geqq \frac{1}{2}(x-y)^T(B-C)(x-y) \geqq 0$$

for each $y \in A(x)$. Moreover, $f(x) = f(y)$ for some $y \in A(x)$ if and only if x solves the LCP(q,M).

The following theorem is the main result if M is symmetric and nondegenerate.

Theorem 2.1. Let M be a symmetric and nondegenerate matrix and let q be an arbitrary vector. Let (B,C) be a regular Q-splitting of M. Let E be a nonnegative diagonal matrix with $E_{ii} < 1$ for all i. Then, the following three statements are equivalent:

(A) For any initial vector $x^0 \geqq 0$, any sequence $\{x^k\}$ satisfying $x^{k+1} \in A(x^k)$ is bounded and thus has at least one accumulation point; moreover, any such point solves the LCP(q,M);

(B) The quadratic function $f(x) = q^Tx + \frac{1}{2}x^TMx$ is bounded below for $x \geqq 0$.

(C) For any initial vector $x^0 \geqq 0$, any sequence $\{x^k\}$ satisfying $x^{k+1} \in A(x^k)$ converges to a solution of the LCP(q,M).

Proof. (A) ==> (B). Let $x^0 \geqq 0$. Use the given x^0 as the initial iterate, generate a sequence $\{x^k\}$ with $x^{k+1} \in A(x^k)$. By (A), some subsequence converges to some solution $\hat{x}$ of the LCP(q,M). By Lemma 2.1, we have

$$f(x^0) \geqq f(\hat{x}) \quad .$$

Since M is nondegenerate, the LCP(q,M) has a finite number of solutions. Consequently, for any $x^0 \geqq 0$, $f(x^0)$ is bounded below by the minimum of the quadratic function values $f(\hat{x})$ generated by a finite set of $\hat{x}$ vectors. Thus (B) follows.

(B) ==> (C). Let $\{x^k\}$ be any sequence satisfying $x^{k+1} \in A(x^k)$ with $x^0 \geqq 0$. According to the proof of Theorem 4.1 in Ref. 1, it suffices to establish two things: (i) that the sequence $\{x^k\}$ is bounded and (ii) that the entire sequence $\{x^k\}$ in fact converges. To prove (i), suppose that the sequence $\{x^k\}$ is unbounded. Then there exist a nonempty index set α and a certain subsequence $\{x^{k_i}\}$ such that $\{x_j^{k_i}\} \to \infty$ if $j \notin \alpha$ and $\{x_j^{k_i}\}$ is bounded if $j \in \alpha$. Let β denote the complement of α.

By Lemma 2.1, the sequence $\{f(x^k)\}$ is non-increasing. Assumption (B) implies that $\{f(x^k)\}$ is bounded below and therefore converges. By Lemma 2.1 again, we have

$$f(x^k) - f(x^{k+1}) \geqq \frac{1}{2} (x^k-x^{k+1})^T(B-C)(x^k-x^{k+1}) \geqq 0 \quad \text{all} \quad k \ .$$

Since B-C is positive definite, it follows that the sequence $\{x^k-x^{k+1}\}$ converges to zero.

Returning to the subsequence $\{x^{k_i}\}$, we have by definition of x^{k_i},

$$q + Cx^{k_i-1} + Bx^{k_i} \geqq 0 \tag{2a}$$

$$x^{k_i} \geqq Ex^{k_i-1} \tag{2b}$$

$$(x^{k_i}-Ex^{k_i-1})^T(q+Cx^{k_i-1}+Bx^{k_i}) = 0 \quad . \tag{2c}$$

Since $\{x_j^{k_i}\} \to \infty$ for $j \in \alpha$ and $\{x^k-x^{k-1}\} \to 0$, it follows that $\{x_j^{k_i-1}\}$ also tends to infinity as $k_i \to \infty$ for $j \in \alpha$. We may write

$$x_j^{k_i} - E_{jj}x_j^{k_i-1} = (x_j^{k_i}-x_j^{k_i-1}) + (1 - E_{jj})x_j^{k_i-1} \quad .$$

Since $E_{jj} < 1$, it follows that

$$x_j^{k_i} > E_{jj}x_j^{k_i-1}$$

for all k_i large enough and all $j \in \alpha$. Hence, by complementarity, we obtain (cf. (2a) - (2c))

$$\begin{aligned} 0 &= (q + Cx^{k_i-1} + Bx^{k_i})_\alpha \\ &= [(q + C(x^{k_i-1} - x^{k_i}))_\alpha + M_{\alpha\beta}x_\beta^{k_i}] + M_{\alpha\alpha}x_\beta^{k_i}] \quad . \end{aligned} \tag{3}$$

which implies by the nondegeneracy of M,

$$x_\alpha^{k_i} = -M_{\alpha\alpha}^{-1}[(q + C(x^{k_i-1} - x^{k_i}))_\alpha + M_{\alpha\beta}x_\beta^{k_i}] \quad .$$

But this is a contradiction because the left-hand side in the above equation is unbounded whereas the right-hand side is bounded. Consequently, the entire sequence $\{x^k\}$ must be bounded. The boundedness of $\{x^k\}$ and the proof of Theorem 4.1 in Ref. 1 imply the following conclusion: The sequence $\{x^k\}$ has at least one accumulation point and any such point solves the LCP(q,M). Since the LCP(q,M) has only a finite number of solutions, the sequence $\{x^k\}$ has a finite number of accumulation points. Thus, by Ostrowski's Theorem which states that a bounded sequence $\{y^k\}$ with a finite number of accumulation points and satisfying $\|y^{k+1}-y^k\| \to 0$ converges (Ref. 4, Theorem 28.1), the sequence $\{x^k\}$ indeed converges. This completes the proof of (B) ==> (C).

(C) ==> (A). This is obvious.

Among other things, Theorem 2.1 has shown that if M is symmetric and nondegenerate, and if assumption (B) holds, then for any initial vector $x^0 \geqq 0$, any sequence $\{x^k\}$ satisfying $x^{k+1} \in A(x)$ is bounded. This statement remains valid if the nondegeneracy assumption of M is replaced by the weaker assumption that the homogeneous LCP(0,M) has zero as the unique solution. To see this, suppose that $\|x^k\| \to \infty$. Then the normalized sequence $\{x^k/\|x^k\|\}$ has an accumulation point $\hat{x}$ which must be nonnegative and nonzero. Let $\{x^{k_i}/\|x^{k_i}\|\}$ be a subsequence converging to $\hat{x}$. For each k_i, we have

$$q + Cx^{k_i-1} + Bx^{k_i} = q + C(x^{k_i-1} - x^{k_i}) + Mx^{k_i} \geqq 0 \quad .$$

The proof of (B) ==> (C) shows that the sequence $\{x^{k_i-1} - x^{k_i}\} \to 0$. Thus, dividing by $\|x^{k_i}\|$ and passing the limit $k_i \to \infty$, we deduce that

$$M\hat{x} \geqq 0 \quad .$$

Moreover, for each index j, we have

$$\begin{aligned} 0 &= (q + Cx^{k_i-1} + Bx^{k_i})_j \, (x^{k_i} - Ex^{k_i-1})_j \\ &= (q + C(x^{k_i-1} - x^{k_i}) + Mx^{k_i})_j \, ((I-E)x^{k_i} - E(x^{k_i-1} - x^{k_i}))_j \quad . \end{aligned}$$

Dividing by $\|x^{k_i}\|^2$ and noting that $E_{jj} < 1$, we obtain by passing the limit $k_i \to \infty$,

$$\hat{x}_j(M\hat{x})_j = 0 \text{ for all } j \quad .$$

Thus $\hat{x}$ is a nonzero solution of the homogeneous LCP(0,M). Consequently, if the homogeneous LCP(0,M) has no nonzero solution, then the sequence $\{x^k\}$ must be bounded.

It is well-known from quadratic programming theory that if a quadratic objective function is bounded below on a feasible region, then it achieves its minimum there (see Refs. 5 & 6). In particular, assumption (B) implies that the quadratic program (1) has a solution and so does the LCP(q,M). Thus, the existence of a solution to the LCP(q,M) is implicit in condition (B). Moreover, (B) holds if and only if the matrix M is copositive (i.e. $x \geqq 0 \Longrightarrow x^T M x \geqq 0$) and the implication below holds:

$$[x \geqq 0,\ x^T M x = 0] \Longrightarrow q^T x \geqq 0 \quad ,$$

see Ref. 6 for a proof. In general, the existence of a solution to the LCP(q,M) does not imply (B).

Two splittings that are particularly interesting are derived from Mangasarian's iterative procedure (Ref. 7) and the block SOR iterative method. Mangasarian's procedure leads to a splitting (B,C) of the form

$$B = K + (\lambda\omega)^{-1}D,\ C = M - K - (\lambda\omega)^{-1}D \ \text{ and } \ E = (1-\lambda)I \tag{4}$$

where $\lambda \in (0,1]$, $\omega > 0$ and D is a positive definite diagonal matrix (see Ref. 1). Specializing Theorem 2.1 to such a splitting, we obtain the following characterization.

<u>Corollary 2.1</u>. Let M be a symmetric and nondegenerate matrix and q an arbitrary vector. Let $\lambda \in (0,1]$ and $\omega > 0$ be given scalars. Suppose that the matrices K and D satisfy:

(i) D is a positive definite diagonal matrix

(ii) the matrix $K + (\lambda\omega)^{-1}D - M/2$ is positive definite

(iii) $K + (\lambda\omega)^{-1}D$ is a Q-matrix.

Then the statements (A), (B) and (C) in Theorem 2.1 are equivalent for Mangasarian's procedure.

<u>Proof</u>. It suffices to observe that the splitting (B,C) defined in (4) is regular and Q under the given assumptions.

Using the same theorem due to Ostrowski, Mangasarian (Ref. 7) has proven the convergence of the sequence of vectors generated by his iterative scheme under the assumption that the matrix M is symmetric, nondegenerate as well as copositive-plus. Note that copositivity is not a pre-stated assumption in Corollary 2.1 (or Theorem 2.1); instead, it is part of the characterizing condition for convergence (implied by condition (B)).

The block SOR-splitting is defined in the following way. Let the matrix M be partitioned into submatrices (M_{ij}) with each diagonal block M_{ii} being square. Let D, L and U be consisted of the diagonal, strictly lower and upper triangular blocks of M respectively. Then the block SOR-splitting (B,C) of M is given by

$$B = L + D/\omega \quad \text{and} \quad C = U + (1 - 1/\omega)D$$

where $\omega \in (0,2)$ is a given scalar. See Ref. 1. The point SOR-splitting corresponds to the case where each M_{ij} is the (i,j) entry of M.

Corollary 2.2. Let M be a symmetric, nondegenerate matrix partitioned into submatrices (M_{ij}) with each diagonal submatrix M_{ii} being positive definite. Let q be an arbitrary vector and E any nonnegative diagonal matrix with $E_{ii} < 1$ all i. Then for all $\omega \in (0,2)$ and any $x^0 \geqq 0$, the sequence $\{x^k\}$ generated by the block SOR-splitting and the matrix E is uniquely defined. Moreover, the statements (A), (B) and (C) in Theorem 2.1 are equivalent for the block SOR-splitting.

Proof. The fact that the sequence $\{x^k\}$ is uniquely defined follows from the positive definiteness of the matrix D which implies that B is a P-matrix. The equivalence of (A), (B) and (C) follows from Theorem 2.1 and the observation that the block SOR-splitting is in fact regular and Q under the assumed properties (see the proof of Corollary 4.2 in Ref. 1).

3. THE POSITIVE SEMI-DEFINITE CASE

The analog of Theorem 2.1 for a positive semi-definite M is the following.

Theorem 3.1. Let M be a symmetric positive semi-definite matrix and let q be an arbitrary vector. Let (B,C) be a regular Q-splitting of M. Let E be a nonnegative diagonal matrix with $E_{ii} < 1$ for all i. Then for any initial vector $x^0 \geqq 0$, the sequence $\{x^k\}$ generated by the splitting (B,C) and the matrix E is uniquely defined. Moreover, the following three statements are equivalent:

(D) For any initial vector $x^0 \geqq 0$, the sequence $\{Mx^k\}$ converges to some vector $M\hat{x}$ and $\hat{x}$ solves the LCP(q,M).

(B) Same as before.

(E) The LCP(q,M) has a solution.

Before proving the theorem, we should point out that statement (D) does not assert even the boundedness of the sequence $\{x^k\}$. In particular, $\{x^k\}$ could be unbounded. However, the sequence $\{Mx^k\}$ must converge and in fact converges in a desirable manner.

Among the equivalence between the statements (D), (B) and (E), the only non-trivial part that requires a detailed proof is the implication (B) (or equivalently (E)) ==> (D). The equivalence of (B) and (E) is well-known by the symmetry and positive semi-definiteness of M. That (D) ==> (E) is also obvious. Condition (B) (or (E)) is further equivalent to the feasibility of the LCP(q,M) which in turn is equivalent to the implication:

$$[v \geqq 0,\ Mv = 0] \Longrightarrow v^T q \geqq 0 \quad .$$

The last implication is satisfied if condition (F) below holds:

(F) There exists a vector v such that $q + Mv > 0$.

By using the same argument as in Lemma 3 of Ref. 7, one can show that

condition (F) (under the assumed properties of M) implies that the sequence $\{x^k\}$ must be bounded and thus assertion (D) follows readily. However, condition (F) is in general stronger than (B) or (E), even for symmetric positive semi-definite M. As a result, the implication (B) ==> (D) requires a separate proof. For this purpose, we need a few lemmas. The first lemma concerns the accumulation point(s) of a sequence in the affine image of a polyhedral set. The second lemma gives a straightforward property of a positive semi-definite matrix. The third lemma is a consequence of the second. Finally, the fourth lemma concerns the solution of an LCP with a positive semi-definite matrix. Although the proof of each of these lemmas is fairly easy, we want to state the lemmas explicitly because they all play an important role in the proof of Theorem 3.1.

Lemma 3.1. Let S be a polyhedral set in R^n and A any $m \times n$ matrix. Then any accumulation point of a sequence $\{Ax^k\}$ where $\{x^k\} \subseteq S$ must be of the form $A\hat{x}$ for some $\hat{x} \in S$.

Proof. This follows easily from the observation that the image AS is also a polyhedral set (see Ref. 8) and is therefore closed.

Lemma 3.2. Let C be an $n \times n$ symmetric positive definite matrix and A any $m \times n$ matrix. Then the sequence $\{ACA^Ty^k\}$ converges to the vector $ACA^T\bar{y}$ if and only if the sequence $\{A^Ty^k\}$ converges to $A^T\bar{y}$.

Proof. It is obvious that if the sequence $\{A^Ty^k\}$ converges to $A^T\bar{y}$, then the sequence $\{ACA^Ty^k\}$ converges to $ACA^T\bar{y}$. To prove the "only if" part, assume that the sequence $\{ACA^Ty^k\}$ converges to $ACA^T\bar{y}$. With no loss of generality, we may take $\bar{y} = 0$. We first show that the sequence $\{A^Ty^k\}$ is bounded. Suppose not. Then $\|A^Ty^k\| \to \infty$. The normalized sequence $\{A^Ty^k/\|A^Ty^k\|\}$ has an accumulation point which must be nonzero and of the form $A^T\hat{y}$ by Lemma 3.1. On the one hand, we have

$$ACA^T y^k / \|A^T y^k\| \to 0$$

because the numerator tends to zero and the denominator tends to ∞. On the other hand, the sequence $\{ACA^T y^k/\|A^T y^k\|\}$ has $ACA^T\hat{y}$ as an accumulation point. Thus, it follows that

$$ACA^T\hat{y} = 0$$

which implies, by the positive definiteness of C,

$$A^T\hat{y} = 0$$

which is a contradiction. Therefore, the sequence $\{A^T y^k\}$ is bounded. Again, by Lemma 3.1, any accumulation point of the sequence $\{A^T y^k\}$ must be of the form $A^T u$ for some vector u. By the same argument just used above, we can easily deduce $A^T u = 0$. Consequently, the sequence $\{A^T y^k\}$ converges to zero as desired.

Remark. According to Lemma 3.1, the sequence $\{ACA^T y^k\}$ converges to some vector of the form $ACA^T\bar{y}$ if and only if it in fact converges.

Lemma 3.3. Let M be a symmetric positive semi-definite matrix. Let α be any index set. Then the sequence $\{M_{\alpha\alpha} y_\alpha^k\}$ converges to the vector $M_{\alpha\alpha}\bar{y}_\alpha$ if and only if the sequence $\{M_{\cdot\alpha} y_\alpha^k\}$ converges to $M_{\cdot\alpha}\bar{y}_\alpha$.

Proof. It suffices to show the "only if" part. Suppose that $\{M_{\alpha\alpha} y_\alpha^k\}$ converges to $M_{\alpha\alpha}\bar{y}_\alpha$. Since M is symmetric positive semi-definite, we may write $M = LL^T$ for some matrix L. Under this representation, we obtain

$$M_{\alpha\alpha} = L_{\alpha\cdot}(L_{\alpha\cdot})^T \quad \text{and} \quad M_{\cdot\alpha} = L(L_{\alpha\cdot})^T \quad .$$

Since $M_{\alpha\alpha} y_\alpha^k = L_{\alpha\cdot}(L_{\alpha\cdot})^T y_\alpha^k$, it follows from Lemma 3.2 that the sequence $\{(L_{\alpha\cdot})^T y_\alpha^k\}$ converges to $(L_{\alpha\cdot})^T\bar{y}_\alpha$ because $M_{\alpha\alpha} y_\alpha^k \to M_{\alpha\alpha}\bar{y}_\alpha$. Consequently, $M_{\cdot\alpha} y_\alpha^k = L(L_{\alpha\cdot})^T y_\alpha^k \to L(L_{\alpha\cdot})^T\bar{y}_\alpha = M_{\cdot\alpha}\bar{y}_\alpha$ as desired.

Lemma 3.4. Let M be a symmetric positive semi-definite matrix and let q be an arbitrary vector. If x^1 and x^2 are two solutions of the LCP(q,M), then $Mx^1 = Mx^2$.

Proof. This is well-known in linear complementarity theory, see Ref. 9 e.g.

Proof of the implication (B) ==> (D) in Theorem 3.1. We first note that if M is positive semi-definite and (B,C) is a regular splitting of M, then B itself must be positive definite. To see this, let $x \neq 0$ be any vector. Then

$$0 \leqq x^T Mx = x^T(B+C)x \quad \text{and} \quad x^T(B-C)x > 0 \quad .$$

Adding the two inequalities gives $x^T Bx > 0$. Thus, the sequence $\{x^k\}$ is uniquely defined.

By the same argument used in the proof of Theorem 2.1, we may deduce that the sequence $\{x^{k+1} - x^k\}$ converges to zero. We claim that the sequence $\{Mx^k\}$ has at least one accumulation point. This is certainly true if $\{x^k\}$ is bounded. Suppose that $\{x^k\}$ is not bounded. Then there exist a nonempty index set α and a subsequence $\{x^{k_i}\}$ so that $\{x_j^{k_i}\} \to \infty$ if $j \in \alpha$ and $\{x_j^{k_i}\}$ is bounded if $j \in \alpha$. Let β be the complement of α. As *in the* proof of Theorem 2.1, we may deduce (cf. (3))

$$M_{\alpha\alpha} x_\alpha^{k_i} = -[(q + C(x^{k_i - 1} - x^{k_i}))_\alpha + M_{\alpha\beta} x_\beta^{k_i}]$$

for all k_i large enough. The last equation shows that the sequence $\{M_{\alpha\alpha} x_\alpha^{k_i}\}$ is bounded. By Lemma 3.3, the sequence $\{M_{\cdot\alpha} x_\alpha^{k_i}\}$ has an accumulation point. Since

$$Mx^{k_i} = M_{\cdot\alpha} x_\alpha^{k_i} + M_{\cdot\beta} x_\beta^{k_i}$$

and $\{x_\beta^{k_i}\}$ is bounded, it follows that the sequence $\{Mx^{k_i}\}$ has an accumulation point. Therefore, so does $\{Mx^k\}$.

Let z be any accumulation point of the sequence $\{Mx^k\}$. Let $\{Mx^{k_i}\}$ be a subsequence converging to z. There exist a possibly empty index set α and a subsequence $\{x^{\ell_i}\}$ of $\{x^{k_i}\}$ so that $\{x_j^{\ell_i}\} \to \infty$ if $j \in \alpha$ and $\{x_j^{\ell_i}\}$ is bounded if $j \notin \alpha$. Let β be the complement of α. Then as before, we have

$$M_{\alpha\alpha}x_\alpha^{\ell_i} = -[(q + C(x^{\ell_i-1} - x^{\ell_i}))_\alpha + M_{\alpha\beta}x_\beta^{\ell_i}] \tag{5}$$

for all ℓ_i large enough. Moreover, the sequence $\{M_{\cdot\alpha}x_\alpha^{\ell_i}\}$ has an accumulation point which must be of the form $M_{\cdot\alpha}\hat{x}_\alpha$ for some $\hat{x}_\alpha \geqq 0$ by Lemma 3.1. With no loss of generality, we may assume that $M_{\cdot\alpha}x_\alpha^{\ell_i} \to M_{\cdot\alpha}\hat{x}_\alpha$. Since the sequence $\{x_\beta^{\ell_i}\}$ is bounded, we may also assume with no loss of generality that $\{x_\beta^{\ell_i}\}$ converges to some vector $\hat{x}_\beta \geqq 0$. It then follows that $z = M\hat{x}$. We claim that $\hat{x}$ solves the LCP(q,M). We have already noted that $\hat{x} \geqq 0$. Passing the limit $\ell_i \to \infty$ in (5), we obtain

$$(q + M\hat{x})_\alpha = 0$$

because $x^{\ell_i-1} - x^{\ell_i} \to 0$. Moreover, for each ℓ_i, we have

$$\begin{aligned} 0 &\leqq (q + Cx^{\ell_i-1} + Bx^{\ell_i})_\beta \\ &= (q + C(x^{\ell_i-1} - x^{\ell_i}) + Mx^{\ell_i})_\beta \end{aligned}$$

and

$$(x_\beta^{\ell_i})^T(q + C(x^{\ell_i-1} - x^{\ell_i}) + Mx^{\ell_i})_\beta = 0 \quad .$$

Passing the limit $\ell_i \to \infty$ and noting that $x_\beta^{\ell_i} \to \hat{x}_\beta$ and $Mx^{\ell_i} \to M\hat{x}$, we conclude that

$$(q + M\hat{x})_\beta \geqq 0 \quad \text{and} \quad \hat{x}_\beta^T(q + M\hat{x})_\beta = 0 \quad .$$

Consequently, $\hat{x}$ solves the LCP(q,M).

Summarizing, we have proven that if z is any accumulation point of the sequence $\{Mx^k\}$, then there exists a solution $\hat{x}$ of the LCP(q,M) such that $z = M\hat{x}$. By Lemma 3.4, there is only one value for such $M\hat{x}$. Consequently, the sequence $\{Mx^k\}$ converges to $M\hat{x}$ where $\hat{x}$ solves the LCP(q,M). This establishes the theorem.

Remark. Although any accumulation point of the sequence $\{Mx^k\}$ must be of the form My for some vector $y \geqq 0$, it is generally not true that any such y will automatically solve the LCP(q,M). What is true is the existence of at least one such y which is a desired solution. This fact is also related to the reason why it is necessary to follow the line of argument used in the latter part of the above proof.

4. APPLICATIONS TO A STRICTLY CONVEX QUADRATIC PROGRAM.

Theorem 3.1 can certainly be specialized to the two important splittings discussed at the end of Section 2. Instead of giving these routine specializations, we give two applications of Theorem 3.1 to the case of a strictly convex program. Consider the quadratic program

$$\text{minimize} \quad c^T x + \frac{1}{2} x^T D x \quad \text{subject to} \quad Ax \geqq b \tag{6}$$

where the matrix D is symmetric positive definite. The Karush-Kuhn-Tucker conditions may be stated as

$$0 = c + Dx - A^T y \tag{7a}$$

$$v = -b + Ax \geqq 0, \; y \geqq 0 \quad \text{and} \quad v^T y = 0 \quad . \tag{7b}$$

Eliminating the x-variables using (7a) and substituting into (7b), we obtain the LCP

$$v = -(b + AD^{-1}c) + AD^{-1}A^T y \geqq 0, \; y \geqq 0 \quad \text{and} \quad v^T y = 0 \tag{8}$$

where the matrix $AD^{-1}A^T$ is symmetric positive semi-definite. Specializing Theorem 3.1 to the above LCP, we obtain

Corollary 4.1. Let D be a symmetric positive definite matrix. Let (B,C) be any Q-regular splitting of the matrix $AD^{-1}A^T$. Let E be any nonnegative diagonal matrix E with $E_{ii} < 1$ for all i. For any initial $y^0 \geqq 0$, the (uniquely defined) sequence $\{y^k\}$ generated by the splitting (B,C) and the matrix E induces a corresponding sequence of iterates $\{x^k\}$ via (7a); namely

$$x^k = D^{-1}(-c + A^T y^k) \quad \text{for all} \quad k \quad .$$

The following two statements are equivalent:

(D)' For any $y^0 \geqq 0$, the induced sequence $\{x^k\}$ converges to the unique solution of the program (6)

(G) The program (6) is feasible, or equivalently, solvable.

Proof. We first observe that the program (6) is feasible if and only if the LCP(8) is so. Thus, according to Theorem 3.1, statement (G) is equivalent to the fact that for any initial $y^0 \geqq 0$, the sequence $\{AD^{-1}A^Ty^k\}$ converges to some vector $AD^{-1}A^T\hat{y}$ and $\hat{y}$ solves the LCP(8). To see that this latter statement is equivalent to (D)', suppose that $\{AD^{-1}A^Ty^k\}$ converges to $AD^{-1}A^T\hat{y}$ where $\hat{y}$ solves the LCP(8). By Lemma 3.2, the sequence $\{A^Ty^k\}$ converges to $A^T\hat{y}$. Thus the induced sequence $\{x^k\}$ converges to the vector

$$\hat{x} = D^{-1}(-c+A^T\hat{y}) \quad . \tag{9}$$

It is obvious that $\hat{x}$ is the unique solution of the program (6). Conversely, suppose that the sequence $\{x^k\}$ which is induced by $\{y^k\}$ for some $y^0 \geqq 0$ converges to the unique solution $\hat{x}$ of the program (6). Then there exists a vector $\hat{y}$ solving the LCP(8) and satisfying (9). It then follows that the sequence $\{A^Ty^k\}$ converges to $A^T\hat{y}$. Therefore, $AD^{-1}A^Ty^k \to AD^{-1}A^T\hat{y}$ as desired. Consequently, the equivalence of (D)' and (G) follows.

Remark. It is in general, not true that if $\hat{x}$ solves the program (6) and $\hat{y}$ is any vector satisfying (9), then $\hat{y}$ solves the LCP(8). Only a vector of multipliers $\hat{y}$ corresponding to the constraints of (6) will both satisfy (9) and solve (8).

The significance of Corollary 4.1 is the following. A useful way to solve the quadratic program (6) by iterative methods is to apply them to the corresponding LCP(8). Under a constraint qualification of the Slater type, one can show that the sequence of vectors $\{y^k\}$ generated is bounded and thus has at least one accumulation point. Moreover, any such point solves the LCP(8) and therefore yields the unique solution to (6) by means of the relation (9). See Ref. 10 e.g. Typically the fact that any accumulation point of the sequence generated solves the LCP is an inherent property of the iterative methods. The proof of the boundedness of the sequence usually

requires a constraint qualification without which the existence of at least one such point is no longer guaranteed. Ref. 11 contains a somewhat more detailed discussion on this important point. Now, Corollary 4.1 says that under absolutely no constraint qualification at all, the very minimal requirement that the program (6) be feasible (or equivalently, solvable) characterizes the convergence of the induced sequence $\{x^k\}$ to the unique global minimum point of (6). Although, not even the boundedness of the sequence $\{y_k\}$ is asserted, the convergence of the induced sequence $\{x^k\}$ should have served all desired practical purposes.

In a similar way, the significance of Theorem 3.1 is also easily seen. Indeed, as we have mentioned earlier, condition (F) which is a Slater constraint qualification implies conclusion (D). Theorem 3.1 says that the very minimal requirement that the LCP(q,M) be feasible (or equivalently, solvable) characterizes the convergence of the sequence of vectors $\{w^k = q+Mx^k\}$ to a desired solution $\hat{w} = q+M\hat{x}$ of the LCP(q,M). Again, the characterization requires no such constraint qualification as (F).

There is yet another approach to apply an iterative method for solving the strictly convex quadratic program (6). This approach was first proposed by Han and Mangasarian (Ref. 12) as a special case of their exact penalty function theory for general nonlinear programs. In what follows, we describe this approach from a different point of view.

The prime motivation for this alternative formulation has to do with the fact that the LCP(8) involves the inverse of the matrix D. For large-scale applications with sparse data, it is not very desirable to invert D because the inversion can easily destroy the sparsity structure. As a result, one is led to investigate the formulation (7) from which (8) was derived. However, (7) is defined by the matrix

$$\begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix}$$

which contains a zero diagonal block. This zero diagonal block prohibits the application of say, the point SOR-method. As a remedy for this, one observes that (7) is equivalent to

$$0 = c + Dx - A^T y$$

$$v = -b + Ax - \gamma A(c+Dx-A^T y) \geqq 0,\ y \geqq 0 \quad \text{and} \quad v^T y = 0$$

which can be rewritten as

$$0 = c + Dx - A^T y \tag{7a}$$

$$v = -b - \gamma Ac - A(\gamma D-I)x + \gamma AA^T y \geqq 0,\ y \geqq 0 \quad \text{and} \quad v^T y = 0 \tag{7b)'$$

where γ is some positive scalar. If A has no vanishing rows, the matrix γAA^T has positive diagonal entries. Thus the point SOR-method is applicable. However, the matrix

$$\begin{pmatrix} D & -A^T \\ -A(\gamma D-I) & \gamma AA^T \end{pmatrix}$$

is nonsymmetric and most likely, not even positive semi-definite. One way to symmetrize the above matrix is to multiply the expression (7a) by the matrix $\gamma D-I$ which will be nonsingular (in fact, positive definite) if $\gamma > 1/\rho$ where ρ is the least eigenvalue of the positive definite matrix D. Thus, if $\gamma > 1/\rho$, problem (7) is equivalent to

$$0 = (\gamma D-I)c + (\gamma D-I)Dx - (\gamma D-I)A^T y \tag{10a}$$

$$v = -b - \gamma Ac - A(\gamma D-I)x + \gamma AA^T y,\ y \geqq 0 \quad \text{and} \quad v^T y = 0 \quad . \tag{10b}$$

This last formulation (10) is precisely the one to which Han and Mangasarian (Ref. 12) proposed the application of the point SOR-method. They have shown

that for $\gamma \geqq 1/\rho$, the matrix

$$M = \begin{pmatrix} (\gamma D-I)D & -(\gamma D-I)A^T \\ -A(\gamma D-I) & \gamma AA^T \end{pmatrix} \tag{11}$$

is symmetric positive semi-definite and that if $\gamma > 1/\rho$, the problem (10) has a solution $(\bar{x}(\gamma), \bar{y}(\gamma))$ such that $\bar{x}(\gamma) = x^*$ where x^* is the unique solution of the program (6), provided that the program (6) is feasible. This latter conclusion also follows easily from our derivation of the problem (10). We remark that if $\gamma > 1/\rho$ and A has linearly independent rows, then the matrix M in (11) is positive definite.

Concerning the convergence of the point SOR-method, Han and Mangasarian (Ref. 12) have shown that if either (i) the matrix A has no vanishing rows and has linearly independent columns, and there exists a vector $\hat{x}$ with $A\hat{x} > b$, or (ii) the matrix A has linearly independent rows, then for $\gamma > 1/\rho$, for any relaxation parameter $\omega \in (0,2)$ and any initial vector (x^0,y^0) with x^0 arbitrary and $y^0 \geqq 0$, the sequence (x^k,y^k) generated is bounded and thus has at least one accumulation point. Moreover, any such point $(\bar{x},\bar{y})$ solves (10). In fact, $\bar{x}$ is the unique optimum solution of (6) to which the sequence $\{x^k\}$ must converge.

The above convergence result of Han and Mangasarian was proven under some linear independence property of the matrix A and in at least one instance, under a constraint qualification as well. The corollary below shows that under absolutely no such restriction on A and no constraint qualification, the same convergence of the sequence $\{x^k\}$ can be asserted. However, as in the two previous results, (Theorem 3.1 and Corollary 4.1), only the convergence of the sequence $\{A^Ty^k\}$, and not the boundedness of $\{y^k\}$, can be proved. Again, for all practical purposes, it is the convergence of the sequence $\{x^k\}$ that is of interest.

Corollary 4.2. Let D be a symmetric positive definite matrix with least eigenvalue $\rho > 0$. Let A be any matrix with no vanishing rows. Fix $\gamma > 1/\rho$. Then for any $\omega \in (0,2)$ and any initial vector (x^0,y^0) with x^0 arbitrary and $y^0 \geqq 0$, the sequence of iterates $\{(x^k,y^k)\}$ generated by the point SOR-method is uniquely defined. Moreover, the following two statements are equivalent:

(H) For any $\omega \in (0,2)$ and any initial vector (x^0,y^0) with x^0 arbitrary and $y^0 \geqq 0$, the sequence $\{(x^k,y^k)\}$ is such that $\{x^k\}$ converges to the unique solution x^* of the program (6) and that $\{A^Ty^k\}$ converges to some vector A^Ty^* where (x^*,y^*) solves the problem (10);

(G) same as before.

Proof. Although the formulation (10) is not exactly the one of a standard LCP, the same analysis of Theorem 3.1 allows one to conclude the equivalence of the following two statements:

(G)' The problem (10) has a solution.

(H)' For any $\omega \in (0,2)$ and any initial vector (x^0,y^0) with x^0 arbitrary and $y^0 \geqq 0$, the sequence $\{Mz^k\}$ where M is given by (11) and $z^k = (x^k,y^k)$ is the sequence generated by the point SOR-method, converges to some vector $M\hat{z}$, where $\hat{z} = (\hat{x},\hat{y})$, and $\hat{z}$ solves (10).

Noting that the matrix $\gamma D-I$ is nonsingular if $\gamma > 1/\rho$, we deduce that problem (10) has a solution if and only if (7), or equivalently (6), has one. Consequently, statements (G) and (G)' are equivalent. Noting that

$$Mz^k = \begin{pmatrix} (\gamma D-I)(Dx^k - A^Ty^k) \\ Ax^k + \gamma A(Dx^k - A^Ty^k) \end{pmatrix} \tag{12}$$

we may easily deduce that (H) ==> (H)'. Conversely, suppose that $Mz^k \to M\hat{z}$ and $\hat{z} = (\hat{x},\hat{y})$ solves (10). Since $\gamma D-I$ is nonsingular, it follows from the

expression (12) that

$$Dx^k - A^T y^k \to D\hat{x} - A^T\hat{y} \quad \text{and} \quad Ax^k \to A\hat{x} \quad . \tag{13}$$

We claim that the sequence $\{x^k\}$ is bounded. Suppose not. Then $\|x^k\| \to \infty$ and the normalized sequence $\{x^k/\|x^k\|\}$ has a nonzero accumulation point u. With no loss of generality, we may assume that u is the limit of the entire sequence $\{x^k/\|x^k\|\}$. It follows from (13) that

$$(Dx^k - A^T y^k)/\|x^k\| \to 0 \quad \text{and} \quad Au = 0 \quad .$$

Since $\{Dx^k/\|x^k\|\}$ is itself converging to Du, the sequence $\{A^T y^k/\|x^k\|\}$ must converge and by Lemma 3.1, it must converge to $A^T v$ for some vector v. Thus, we have

$$Du - A^T v = 0 \quad .$$

Pre-multiplying the above identity by u^T and using the fact that $Au = 0$, we deduce $u^T Du = 0$ which is impossible because D is positive definite and u is nonzero. Consequently, the sequence $\{x^k\}$ is bounded. Let $\tilde{x}$ be an accumulation point of $\{x^k\}$. We claim that $\tilde{x} = \hat{x}$. Indeed, let $\{x^{k_i}\}$ be a subsequence converging to $\tilde{x}$. Since the sequence $\{Dx^k - A^T y^k\}$ is convergent, it follows that the subsequence $\{A^T y^{k_i}\}$ must itself be converging and by Lemma 3.1, to a vector of the form $A^T v$. Thus, we have

$$D\tilde{x} - A^T v = D\hat{x} - A^T\hat{y} \quad \text{and} \quad A\tilde{x} = A\hat{x}$$

or equivalently.

$$\begin{pmatrix} D & -A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} \tilde{x} - \hat{x} \\ v - \hat{y} \end{pmatrix} = 0$$

which implies, by the positive definiteness of D,

$$\tilde{x} = \hat{x} \quad .$$

Consequently, the sequence $\{x^k\}$ converges to $\hat{x}$ and the sequence $\{A^T y^k\}$ converges to $A^T\hat{y}$. Since $(\hat{x},\hat{y})$ solves (10), it follows that $\hat{x}$ is the unique solution of the quadratic program (6). Therefore, assertion (H) follows. This completes the proof of the corollary.

Remark. At the point where we have established the boundedness of the sequence $\{x^k\}$, we cannot use the Han and Mangasarian convergence result to conclude that $\{x^k\}$ converges to the unique solution of (6). This is because for the Han and Mangasarian result to be applicable, we need to know the boundedness of the sequence $\{y^k\}$ as well. However, the boundedness of $\{y^k\}$ is generally not guaranteed without the linear independence property of the matrix A.

5. ON LAGRANGIAN RELAXATION

In a recent report (Ref. 13), Cottle and Duvall have described a Lagrangian relaxation algorithm for a constrained matrix problem which is formulated as a strictly convex separable quadratic program. They proved the convergence of their algorithm under a strong consistency condition which is a constraint qualification of the Slater type. In what follows, we use the previous analysis to derive a convergence result for the Lagrangian relaxation method under no such constraint qualification. As a matter of fact, we shall treat the following more general strictly convex quadratic program

$$\text{minimize } \frac{1}{2} x^T D x + c^T x \quad \text{subject to} \quad Ax \geqq b \quad \text{and} \quad x \in X \tag{14}$$

where the matrix D is symmetric positive definite and the set X is polyhedral:

$$X = \{x \in R^n : Fx \geqq f\} \quad .$$

As with the previous problem (6), we have included only inequality constraints in the program (14). This is done for the sake of consistency; the same analysis is equally applicable to programs with both inequality and equality constraints (such as the constrained matrix problem in Ref. 13).

The Lagrangian dual function of the program (14) is

$$d(y) = \min_{x \in X} \frac{1}{2} x^T D x + c^T x + y^T(b - A^T x) \tag{15a}$$

and the dual of (14) can be defined as

$$\max_{y \geqq 0} d(y) \quad . \tag{15b}$$

The Lagrangian relaxation approach for solving the program (14) may be described as follows. Let

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_N \end{pmatrix}$$

be a partition of the rows of the matrix A. Let $y^0 = (y_1^0,\ldots,y_N^0)$ be an arbitrary nonnegative vector partitioned in accordance with A. In general, let $y^k = (y_1^k,\ldots,y_N^k) \geqq 0$ be given. To obtain $y^{k+1} = (y_1^{k+1},\ldots,y_N^{k+1})$, solve N quadratic subprograms consecutively for $\alpha = 1,\ldots,N$:

$$\begin{aligned} &\text{minimize } \frac{1}{2} x^T D x + (c^T - \sum_{\beta<\alpha} (y_\beta^{k+1})^T A_\beta - \sum_{\beta>\alpha} (y_\beta^k)^T A_\beta) x \\ &\text{subject to} \qquad A_\alpha x \geqq b_\alpha \text{ and } x \in X \end{aligned} \tag{16$_\alpha$}$$

and let $y_\alpha^{k+1} \geqq 0$ denote the vector of Lagrange multipliers for the constraints $(A_\alpha x \geqq b_\alpha)$.

In order for the above Lagrangian relaxation method to be practically useful, it is important that the partitioning of the matrix A should be such that each subproblem $(16)_\alpha$ is very easy to solve. This is indeed the case for the constrained matrix problem, see Ref. 13.

As noted by Cottle and Duvall (Ref. 13), the above Lagrangian relaxation algorithm can be interpreted as a block cyclic ascent method applied to the dual program (15). Indeed, it is easy to see that the vector y_α^{k+1} solves the dual subprogram

$$\max_{y_\alpha \geqq 0} d(\{y_\beta^{k+1}\}_{\beta<\alpha}, y_\alpha, \{y_\beta^k\}_{\beta>\alpha})$$

which is obtained from (15b) by fixing all variables except those of y_α.

In order to state our convergence result for the Lagrangian relaxation method, we give a different interpretation of the method in the context of matrix splittings. To simplify the notations, we take $N = 2$ and consider the matrix A partitioned into two sets of rows A_1 and A_2. Then, using the system of linear inequalities which define the set X, we may state the Karush-Kuhn-Tucker conditions of the program (14) as:

$$0 = c + Dx - A_1^T y_1 - A_2^T y_2 - F^T z \qquad x \text{ unrestricted} \tag{17a}$$

$$w_1 = -b_1 + A_1 x \geqq 0 , \quad y_1 \geqq 0 , \quad w_1^T y_1 = 0 \tag{17b}$$

$$w_2 = -b_2 + A_2 x \geqq 0 , \quad y_2 \geqq 0 , \quad w_2^T y_2 = 0 \tag{17c}$$

$$v = -f + Fx \geqq 0 , \quad z \geqq 0 , \quad v^T z = 0 . \tag{17d}$$

Solving for x in (17a) and substituting into (17b) - (17d), we obtain an LCP(q,M) where

$$q = -\begin{pmatrix} b_1 \\ b_2 \\ f \end{pmatrix} - \begin{pmatrix} A_1 \\ A_2 \\ F \end{pmatrix} D^{-1} c \quad \text{and} \quad M = \begin{pmatrix} A_1 \\ A_2 \\ F \end{pmatrix} D^{-1} (A_1^T \; A_2^T \; F^T) \quad . \tag{18}$$

It is then easy to see that given $y^k = (y_1^k, y_2^k) \geqq 0$ and $z^k \geqq 0$, the next iterate $y^{k+1} = (y_1^{k+1}, y_2^{k+1})$ produced by the Lagrangian relaxation algorithm can also be obtained in the following way:

(i) Solve the LCP$(q + C^k u^k, B^k)$ where

$$u^k = \begin{pmatrix} y_1^k \\ y_2^k \\ z^k \end{pmatrix} , \quad B^k = \begin{pmatrix} A_1 D^{-1} A_1^T & 0 & A_1 D^{-1} F^T \\ A_2 D^{-1} A_1^T & A_2 D^{-1} A_2^T & A_2 D^{-1} F^T \\ F D^{-1} A_1^T & 0 & F D^{-1} F^T \end{pmatrix} \quad C^k = M - B^k$$

and let

$$u^{k+\frac{1}{2}} = \begin{pmatrix} y_1^{k+1} \\ y_2^{k+\frac{1}{2}} \\ z^{k+\frac{1}{2}} \end{pmatrix}$$

be a solution;

(ii) Solve the $LCP(q + C^{k+\frac{1}{2}} u^{k+\frac{1}{2}}, B^{k+\frac{1}{2}})$ where

$$B^{k+\frac{1}{2}} = \begin{pmatrix} A_1D^{-1}A_1^T & A_1D^{-1}A_2^T & A_1D^{-1}F^T \\ 0 & A_2D^{-1}A_2^T & A_2D^{-1}F^T \\ 0 & FD^{-1}A_2^T & FD^{-1}F^T \end{pmatrix} \quad C^{k+\frac{1}{2}} = M - B^{k+\frac{1}{2}} ,$$

and let

$$u^{k+1} = \begin{pmatrix} \tilde{y}_1^{k+1} \\ y_2^{k+1} \\ z^{k+1} \end{pmatrix}$$

be a solution.

Observe that in step (i) above, the subvector $y_2^{k+\frac{1}{2}}$ is actually useless for all subsequent computations, thus it can be ignored. Moreover, $(y_1^{k+1}, z^{k+\frac{1}{2}})$ is a solution to the $LCP(\tilde{q}^k, \tilde{B}^k)$ where

$$\tilde{q}^k = -\begin{pmatrix} b_1 \\ f \end{pmatrix} - \begin{pmatrix} A_1 \\ F \end{pmatrix} D^{-1}(c - A_2^T y_2^k) \quad \text{and} \quad \tilde{B}^k = \begin{pmatrix} A_1 \\ F_1 \end{pmatrix} D^{-1}(A_1^T, F_1^T) \quad .$$

The latter LCP is clearly equivalent to the subprogram $(16)_1$. Similarly, step (ii) is equivalent to the subprogram $(16)_2$ and the subvector $\tilde{y}_1^{k+1}$ in u^{k+1} can be ignored in the computations.

To summarize the above discussion, we conclude that the Lagrangian relaxation algorithm can be interpreted as a special realization of our basic iterative scheme for solving the LCP in which the splitting (B^k,C^k) is changing from one iteration to another in a cyclic order. With this interpretation on hand, we can now state and prove our convergence result for the Lagrangian method under no constraint qualification on the original program (14).

Theorem 5.1. Let D be a symmetric positive definite matrix. Suppose that each of the subprograms $(16)_\alpha$ is feasible. Let $\tilde{x}(\hat{y}^{k,\alpha})$ denote the unique optimum solution of $(16)_\alpha$ where

$$\hat{y}^{k,\alpha} = (y_1^{k+1},\ldots,y_{\alpha-1}^{k+1}, y_\alpha^k,\ldots,y_N^k) \quad .$$

Then the following two statements are equivalent:

(I) For any $y^0 = (y_\alpha^0) \geqq 0$, the sequence $\{\tilde{x}(\hat{y}^{k,\alpha})\}$ converges to the unique global minimizer of the program (14);

(J) The program (14) is feasible, or equivalently, solvable.

Proof. It suffices to show (J) ==> (I). Suppose that (14) has an optimum solution x^* which must necessarily be unique. Let $\{\tilde{x}(\hat{y}^{k,\alpha})\}$ be the sequence generated by some initial vector $y^0 = (y_\alpha^0) \geqq 0$. We sketch the proof that $\{\tilde{x}(\hat{y}^{k,\alpha})\} \to x^*$ for the case $N = 2$. The argument for an arbitrary N is similar. According to the matrix-splitting interpretation of the Lagrangian relaxation method, we see that the vector y^{k+1} is obtained from y^k by solving two LCP's, namely $(q + C^k u^k, B^k)$ and $(q + C^{k+\frac{1}{2}} u^{k+\frac{1}{2}}, B^{k+\frac{1}{2}})$.

The proof that $\{\tilde{x}(\hat{y}^{k,\alpha})\}$ converges to x^* is essentially the same as that of Corollary 4.1 which is a consequence of Theorem 3.1. However, a key requirement, namely, that $\{u^{k+1} - u^k\} \to 0$ in the proof of the implication (B) ==> (D) in Theorem 3.1 is not guaranteed in the present situation. This is because the splittings (B^k,C^k) and $(B^{k+\frac{1}{2}},C^{k+\frac{1}{2}})$ are not necessarily

regular. In fact, it can easily be shown that the matrices $B^k - C^k$ and $B^{k+\frac{1}{2}} - C^{k+\frac{1}{2}}$ are positive semi-definite but need not be positive definite. A closer look at the proof of (B) ==> (D) in Theorem 3.1 reveals that if

$$C^k(u^{k+\frac{1}{2}} - u^k) \to 0 \quad \text{and} \quad C^{k+\frac{1}{2}}(u^{k+1} - u^{k+\frac{1}{2}}) \to 0 , \tag{19}$$

then the conclusion that the sequence $\{Mu^k\}$ (and also the sequence $\{Mu^{k+\frac{1}{2}}\}$) converges to some $M\hat{u}$ where $\hat{u}$ solves the LCP(q,M) remains valid. Thus, if (19) holds, the convergence of $\{\tilde{x}(\hat{y}^{k,\alpha})\}$ to x^* follows as in Corollary 4.1. The proof that (19) holds resembles that of $\{x^{k+1} - x^k\} \to 0$ in Theorem 3.1 and is not repeated here.

Toward the end of the report (Ref. 13), Cottle and Duvall presented some computational experience to support the superiority of a modified version of the above Lagrangian method. This modified version involves the relaxation of each iterate y_α^{k+1} as in a block SOR method. At present time, we are not able to prove the convergence of this modified method. The main difficulty is that we can not prove the key descent property (cf. Lemma 2.1) when the iterates y_α^{k+1} are overrelaxed. This difficulty can be partially explained by considering the matrix-splitting interpretation of the Lagrangian method. Consider the case $N = 2$. The vectors y_1^{k+1} and y_2^{k+1} appear as subvectors in $u^{k+\frac{1}{2}}$ and u^{k+1}. By overrelaxing only y_1^{k+1} and y_2^{k+2}, and not the entire vectors $u^{k+\frac{1}{2}}$ and u^{k+1}, it is not clear whether the quadratic function

$$\theta(u) = q^T u + \frac{1}{2} u^T M u$$

where q and M are given in (18) will have the desired descent property. After all, this descent property has played such an indispensable role in all the convergence results obtained here.

6. ON GRADIENT-TYPE METHODS

In the previous two sections, we have studied the convergence of three iterative approaches for solving a strictly convex quadratic program. All three approaches are based on a dual formulation of the original program. There is a fourth approach which is also based on a dual formulation and which makes use of a gradient-type method for unconstrained optimization. This latter approach was proposed by Ha (Ref. 14) for solving structured, large-scale quadratic programs. To describe Ha's approach, consider the quadratic program

$$\text{minimize } \frac{1}{2} x^T Dx + c^T x \quad \text{subject to} \quad Ax = b \quad \text{and} \quad x \in X \tag{20}$$

where D is a symmetric positive definite matrix and X is a polyhedral set. (In the case where D is positive semi-definite, Ha has proposed the use of the proximal point algorithm (Ref. 15) to strongly convexify the objective function. Much more detailed discussions on how the proximal point algorithm can be used as a decomposition method for solving large-scale structured convex programs can be found in Ha's Ph.D. dissertation (Ref. 16).) The dual of (20) may be stated as

$$\text{maximize } d(y) : y \text{ unrestricted} \tag{21}$$

where $d(y)$ is as given by (15a). If the set X is nonempty, then the dual function is finite, concave and differentiable for all y. However, $d(y)$ is typically not strongly concave. The gradient of $d(y)$ is given by

$$\nabla d(y) = b - A\tilde{x}(y) \tag{22}$$

where $\tilde{x}(y)$ is the unique global minimizer of the Lagrangian function

$$L(x,y) = \frac{1}{2} x^T Dx + c^T x + y^T(b-Ax)$$

over $x \in X$.

Problem (21) is an unconstrained maximization program with a concave everywhere differentiable objective function. (The unrestrictedness of the y-variables is due to the equality constraints $Ax = b$. In general, given a strictly convex quadratic program such as (14) where there are inequality constraints, the process of converting such inequality constraints to equality constraints by adding or subtracting slack variables will destroy the strict (equivalent to strong in the quadratic case) convexity of the objective function; thus a strong convexification procedure might be necessary in such a case.) Ha (Ref. 14) proposed the use of a gradient-type method for unconstrained optimization to solve (21). A point of concern here is the boundedness of the sequence generated. Typically, such a boundedness conclusion follows as an immediate consequence of a boundedness assumption on the level sets of the function to be optimized. In the case of problem (21), we have the following characterization.

Proposition 6.1. Let D be a symmetric positive definite matrix and let X be a nonempty polyhedral set. Then the dual function $d(y)$ defined in (15a) has bounded level sets if and only if there exists a neighborhood U of b such that the system

$$Ax = b'$$

has a solution in X for all $b' \in U$. In particular, if $d(y)$ has bounded level sets, then A must have linearly independent rows.

Proof. According to Corollary 14.2.2 in Ref. 8, the function $d(y)$ has bounded level sets if and only if $0 \in \text{int}(\text{dom } d^*)$ where $\text{dom } d^*$ denotes the effective domain of the conjugate function d^* of d. By an easy manipulation, it can be shown that

$$\text{dom } d^* = \{z : \text{the system } Ax = b - z,\ x \in X \text{ is consistent}\} \ .$$

Thus $0 \in \text{int}(\text{dom } d^*)$ if and only if there exists a neighborhood N of the

origin such that for all vectors $z \in N$ the system $Ax = b - z$ has a solution $x \in X$. By taking $U = b + N$, we deduce the desired characterization. To show the last statement, suppose that $d(y)$ has bounded level sets but there exists a nonzero vector u such that $u^T A = 0$. Choose a vector v so that $u^T v \neq 0$ and $b \pm v \in U$. Then there exist vectors x and x' in in X with

$$Ax = b + v \quad \text{and} \quad Ax' = b - v \quad .$$

Pre-multiplying the two expressions by u^T and subtracting, we deduce $u^T v = 0$ which is a contradiction. Therefore, the matrix A must have linearly independent rows.

From the above Proposition, we see that the boundedness of the level sets of the dual function $d(y)$ implies a kind of stability property on the constraint system of the program (20). The following theorem shows that regardless of whether this stability property holds, a sequence of primal vectors $\{\tilde{x}(y^k)\}$ induced by a sequence $\{y^k\}$ which is generated by a broad class of gradient-type methods applied to the dual program (21) will always converge to unique global minimizer of (20), provided that the minimizer exists and that X is either a polyhedral cone or a bounded polyhedral set.

<u>Theorem 6.1</u>. Let D be a symmetric positive definite matrix and let X be either a polyhedral cone or a bounded polyhedral set. Let $\{y^k\}$ be a sequence of dual variables which induces a sequence of primal variables $\{\tilde{x}(y^k)\}$ where

$$\tilde{x}(y) = \arg\min_{x \in X} \frac{1}{2} x^T D x + c^T x + y^T (b - Ax) \quad .$$

Then the following two statements are equivalent:

(K) The sequence $\{\tilde{x}(y^k)\}$ converges to the unique global minimizer x^* of (20);

(L) The sequence $\{\nabla d(y^k)\}$ converges to zero.

Theorem 6.1 has captured the essence of the convergence results in Corollaries 4.1 and 4.2 and Theorem 5.1 by giving a necessary and sufficient condition on the dual sequence $\{y^k\}$ in order for the corresponding primal sequence $\{\tilde{x}(y^k)\}$ to converge to the unique solution of (20). To prove Theorem 6.1, we need the following lemma which summarizes several important properties of $\tilde{x}(y)$ considered as a function of y.

<u>Lemma 6.1</u>. Let D be a symmetric positive matrix of order n and let X be a nonempty convex polyhedral set in R^n. Define $x^* : R^n \to R^n$ by

$$x^*(q) = \arg\min_{x \in X} \frac{1}{2} x^T D x + q^T x \ .$$

Then

(i) $x^*(\cdot)$ is a well-defined Lipschitz continuous function of q, and there exists a positive constant K such that

$$\|x^*(q^1) - x^*(q^2)\|^2 \leq -K(q^1-q^2)^T(x^*(q^1)-x^*(q^2)) \quad \text{for all } q^1,q^2 \ . \tag{23}$$

In particular, if $\{x^*(q^k)\}$ and $\{x^*(p^k)\}$ are two sequences such that $q^k - p^k \to 0$, then $x^*(q^k) - x^*(p^k) \to 0$.

(ii) If Q is convex polyhedral set, then the image $x^*(Q)$ is closed. In particular, if $\{x^*(q^k)\}$ is a converging sequence with $\{q^k\} \subseteq Q$, then there exists a vector $\hat{q} \in Q$ such that $x^*(q^k) \to x^*(\hat{q})$.

(iii) If X is a polyhedral cone, then

$$x^*(\alpha q) = \alpha x^*(q) \quad \text{for all } \alpha \geq 0 \ . \tag{24}$$

<u>Proof</u>. (i) The well-definedness of $x^*(\cdot)$ is clear. To show (23), we note that

$$(x-x^*(q))^T(Dx^*(q)+q) \geq 0 \quad \text{for all } x \in X \tag{25}$$

by the variational inequality characterization of $x^*(q)$. Thus, it follows that

$$(x^*(q^1)-x^*(q^2))^T(Dx^*(q^2)+q^2) \geq 0 \quad \text{and} \quad (x^*(q^2)-x^*(q^1))^T(Dx^*(q^1)+q^1) \geq 0 \ .$$

Adding the two inequalities gives

$$(x^*(q^1)-x^*(q^2))^T D(x^*(q^1)-x^*(q^2)) \leq -(q^1-q^2)^T(x^*(q^1)-x^*(q^2)) \ .$$

By taking K to be the reciprocal of the smallest eigenvalue of D, we obtain (23). From (23), we deduce

$$\|x^*(q^1)-x^*(q^2)\|_2 \leq K\|q^1-q^2\|_2 \quad \text{for all} \quad q^1,q^2$$

which establishes the Lipschitz continuity of $x^*(q)$ as well as the last conclusion of (i).

(ii) If Q is a convex polyhedral set, then $x^*(Q)$ is the union of a finite number of polyhedral sets (see Ref. 17) and is therefore closed. The assertion about the sequence $\{x^*(q^k)\}$ is immediate.

(iii) Obviously, (24) holds for $\alpha = 0$. Let $\alpha > 0$. Multiplying (25) by the scalar α^2, we obtain

$$(\alpha x - \alpha x^*(q))^T(D(\alpha x^*(q)) + \alpha q) \geq 0 \quad \text{for all} \quad x \in X \ .$$

Since X is a cone, $\alpha X = X$, thus by the uniqueness of $x^*(\alpha q)$, we conclude that (24) holds.

Proof of Theorem 6.1. (K) ==> (L). By (22), we have

$$\nabla d(y^k) = b - A\tilde{x}(y^k) \to b - Ax^* = 0$$

because x^* solves (20).

(L) ==> (K). We first prove that the sequence $\{\tilde{x}(y^k)\}$ is bounded. This is certainly true if the set X is bounded. So, let X be a polyhedral cone. Suppose that $\{\tilde{x}(y^k)\}$ is unbounded. Then $\|\tilde{x}(y^k)\| \to \infty$. The normalized sequence $\{\tilde{x}(y^k)/\|\tilde{x}(y^k)\|\}$ has an accumulation point which must be nonzero. Let u be one such point and with no loss of generality, we may assume that $\tilde{x}(y^k)/\|\tilde{x}(y^k)\| \to u$. Since $\nabla d(y^k) = b - A\tilde{x}(y^k) \to 0$, we deduce $Au = 0$. By Lemma 6.1 (iii), we may write

$$\frac{\tilde{x}(y^k)}{\|\tilde{x}(y^k)\|} = \frac{x^*(c-A^Ty^k)}{\|\tilde{x}(y^k)\|} = x^*((c-A^Ty^k)/\|\tilde{x}(y^k)\|) \ .$$

By Lemma 6.1(i), the sequence $\{x^*(-A^T y^k/\|\tilde{x}(y^k)\|) \to u$. Since $A^T(R^n)$ is a convex polyhedral set, Lemma 6.1(ii) implies the existence of a vector $\hat{y}$ such that $u = x^*(A^T\hat{y})$. Hence $Ax^*(A^T\hat{y}) = 0$. By letting $q^1 = A^T\hat{y}$ and $q^2 = 0$, we deduce from (23) that $u = x^*(A^T y) = 0$ which is a contradiction. Consequently, the sequence $\{\tilde{x}(y^k)\}$ is bounded. Let $\hat{x}$ be any accumulation point of $\{\tilde{x}(y^k)\}$. By Lemma 6.1(ii) again, there exists a vector $\hat{y}$ such that $\hat{x} = \tilde{x}(\hat{y})$. Obviously $\hat{x} \in X$. Since $\nabla d(y^k) = b - A\tilde{x}(y^k) \to 0$, it follows that $b - A\hat{x} = 0$. Thus the vector $\hat{x}$ is feasible to (20). Moreover, we have

$$\frac{1}{2}\hat{x}D\hat{x} + c^T\hat{x} = \frac{1}{2}\hat{x}D\hat{x} + c^T\hat{x} + \hat{y}(b-A\hat{x}) = d(\hat{y})$$

because $\hat{x} = \tilde{x}(\hat{y}) = \arg\min_{x \in X} \frac{1}{2}x^T Dx + c^T x + \hat{y}^T(b-Ax)$. Therefore $\hat{x}$ solves the primal program (20). But since (20) has a unique minimizer, say x^*, the sequence $\{\tilde{x}(y^k)\}$ must converge to x^*. This completes the proof of the theorem.

We should point out that condition (L) holds for a large class of gradient-type ascent methods which include many best-known quasi-Newton methods with suitably chosen steplength rules, see Refs. 18 and 19 e.g. According to Theorem 6.1, these gradient methods will all produce a primal sequence converging to the unique solution of the original program (20). Typically, the primal sequence $\{\tilde{x}(y^k)\}$ is generated while the dual sequence $\{y^k\}$ is being computed.

Basically, the proof of Theorem 6.1 shows that assertion (L) ==> (K) requires no particular assumption on the set X. On the other hand, the reverse implication (K) ==> (L) requires that X be either a cone or bounded. At present time, we cannot prove (K) ==> (L) without such assumption on X. We leave this as an open conjecture and encourage the readers to prove or disprove it.

REFERENCES

1. PANG, J. S., Necessary and Sufficient Conditions for the Convergence of Iterative Methods for the Linear Complementarity Problem, to appear in Journal of Optimization Theory and Applications.

2. PANG, J. S., On the Convergence of a Basic Iterative Method for the Implicit Complementarity Problem, Journal of Optimization Theory and Applications, Vol. 17, pp. 149-162, 1982.

3. MURTY, K. G., On the Number of Solutions to the Complementarity Problem and Spanning Properties of Complementarity Cones, Linear Algebra and its Applications, Vol. 5, pp. 65-108, 1972.

4. OSTROWSKI, A. M., Solution of Equations and Systems of Equations, Second Edition, Academic Press, New York, 1966.

5. FRANK, M. and WOLFE, P., An Algorithm for Quadratic Programming, Naval Research Logistics Quarterly, Vol. 3, pp. 95-110, 1956.

6. EAVES, B. C., On Quadratic Programming, Management Science, Vol. 17, pp. 698-711, 1971.

7. MANGASARIAN, O. L., Solution of Symmetric Linear Complementarity Problems by Iterative Methods, Journal of Optimization Theory and Applications, Vol. 22, pp. 465-485, 1977.

8. ROCKAFELLAR, R. T., Convex Analysis, Princeton University Press, New Jersey, 1972.

9. EAVES, B. C., The Linear Complementarity Problem, Management Science, Vol. 17, pp. 68-75, 1971.

10. MANGASARIAN, O. L., Sparsity-Preserving SOR Algorithms for Separable Quadratic and Linear Programming Problems, to appear in Computers and Operations Research.

11. COTTLE, R. W. and PANG, J. S., On the Convergence of a Block Successive Overrelaxation Methods for a Class of Linear Complementarity Problems, Mathematical Programming Study, Vol. 17, pp. 126-138, 1982.

12. HAN, S. P. and MANGASARIAN, O. L., A Dual Differentiable Exact Penalty Function, Mathematical Programming, Vol. 25, pp. 293-306, 1983.

13. COTTLE, R. W. and DUVALL, S. G., A Lagrangian Relaxation Algorithm for the Constrained Matrix Problem, Technical Report SOL 82-10, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, July 1982.

14. HA, C. D., An Algorithm for Structured, Large-Scale Quadratic Programming Problems, Technical Report #2276, Mathematics Research Center, University of Wisconsin-Madison, Sept. 1981.

15. ROCKAFELLAR, R. T., Augmented Lagrangians and Applications of the Proximal Point Algorithm in Convex Programming, Mathematics of Operations Research, Vol. 1, pp. 97-116, 1976.

16. HA, C. D., Decomposition Methods for Structured Convex Programming, Ph.D. dissertation, Dept. of Industrial Engineering, University of Wisconsin-Madison, 1980.

17. ROBINSON, S. M., Some Continuity Properties of Polyhedral Multifunctions, Mathematical Programming Study, Vol. 14, pp. 206-214, 1981.

18. ORTEGA, J. M. and RHEINBOLDT, W. C., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York 1970.

19. DENNIS, J. E., JR. and MORÉ, J. J., Quasi-Newton Methods, Motivation and Theory, SIAM Review, Vol. 19, pp. 46-89, 1977.

JSP/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2552	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) More Results on the Convergence of Iterative Methods for the Symmetric Linear Complementarity Problem		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Jong-Shi Pang		8. CONTRACT OR GRANT NUMBER(s) ECS-8114571 DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Optimization and Large Scale Systems
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE August 1983
		13. NUMBER OF PAGES 37
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, DC 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Convergence, iterative methods, linear complementarity problem, quadratic programming, matrix splitting

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In an earlier paper, the author has given some necessary and sufficient conditions for the convergence of iterative methods for solving the linear complementarity problem. These conditions may be viewed as global in the sense that they apply to the methods regardless of the constant vector in the linear complementarity problem. More precisely, the conditions characterize a certain class of matrices for which the iterative methods will converge, in a certain sense, to a solution of the linear complementarity problem for all constant vectors. In this paper, we improve on our previous results and establish

(continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT (continued)

necessary and sufficient conditions for the convergence of iterative methods for solving each individual linear complementarity problem with a fixed constant vector. Unlike the earlier paper, our present analysis applies only to the symmetric linear complementarity problem. Various applications to a strictly convex quadratic program are also given.

DAT

FILM